

Cite this article: Sanchan, N. (2024). Comparative study on automated reference summary generation using BERT Models and ROUGE score assessment. *Journal of Current Science and Technology*, 14(2), Article 26. <https://doi.org/10.59796/jest.V14N2.2024.26>



## Comparative Study on Automated Reference Summary Generation using BERT Models and ROUGE Score Assessment

Nattapong Sanchan

School of Information Technology and Innovation, Bangkok University, Pathum Thani 12120, Thailand

E-mail: [nattapong.sa@bu.ac.th](mailto:nattapong.sa@bu.ac.th)

Received 6 October, 2023; Revised 31 October, 2023; Accepted 15 December, 2023  
Published online 2 May, 2024

### Abstract

Automatic text summarization is a sub-area in text mining in which a computer system determines the most informative information in the original text to produce a summary for certain jobs and users. In the development of the systems, one of the most important tasks is to evaluate the quality of summaries produced by the systems. Generally, the evaluation task becomes laborious, time-consuming, and expensive because it requires significant efforts on annotation tasks for humans to manually create reference summaries. Being able to generate automatic reference summaries would promote the development of summarization systems in term of speed and evaluation. In this paper, we proposed an Auto-Ref Summary Generation framework for automatically generating reference summaries used in the generic text summarization evaluation task, the Sliced Summary. Given a set of clusters from a cluster ground-truth label dataset, variants of BERT models were utilized for creating cluster representations. The automatic reference summaries were later generated through a centroid-based summarization approach. Overall, DistilBERT, ROBERTa, and SBERT have played crucial roles in automatic summary generation, achieving the highest ROUGE-1 score of 0.47060. However, this does not meet our expectation on text coherence and readability aspects. Although the summaries generated through our proposed framework could not be used as the replacement of the manual summaries, this study has shed new light on the acquisition of automatic reference summaries from a ground-truth label dataset.

**Keywords:** automatic summarization; document clustering; k-means; centroid-based summarization; natural language processing; text mining

### 1. Introduction

Owing to the exceptional growth of textual information in online media, automatic text summarization has emerged to facilitate users to rapidly digest content in text. Automatic text summarization is a sub-area of text mining in which a system determines the most informative information in the original text to produce a summary for certain tasks and users. To generate a summary, researchers have developed summarization systems for different purposes (i.e., single document summarization (Liu et

al., 2019b), multi-document summarization (Chen et al., 2023), aspect-based opinion summarization (Wu et al., 2016), query-focused summarization (Baumel et al., 2016), update summarization (Delort, & Alfonseca, 2012), and cross-language document summarization (Wan, 2011) to summarize different text genres such as product reviews (Yu et al., 2016), news articles (Huang et al., 2011), political text (Sharevski et al., 2021), meeting text (Oya et al., 2014), scientific articles (Altmami, & Menai, 2022), online debates (Sanchan et al., 2017; Sanchan,